

## RP 505 Asset Identification and Data Reporting

The following recommended practice (RP) is subject to the disclaimer at the front of this manual. It is important that users read the disclaimer before considering adoption of any portion of this recommended practice.

This recommended practice was prepared by a committee of the AWEA Operations and Maintenance (O&M) Committee.

Committee Chair: Bruce Hamilton, Navigant  
Principal Author: David Zeglinski, OSIsoft, LLC  
Contributing Author:  
Alistair Ogilvie, Sandia National Laboratory  
Ben Karlson, Sandia National Laboratory

### **Purpose and Scope**

The scope of “Asset Identification and Data Reporting” focuses on generating, collecting, and serving up wind farm data for asset management and maintenance and reliability improvements. There are two key components: the metadata system associated with the generating assets and the data collection system itself.

### **Data Reporting**

#### **1. Metadata System**

##### **1.1. Develop a Detailed Taxonomy**

Developing a detailed equipment breakdown, or taxonomy, helps ensure that maintenance data is captured with enough detail to be useful. Using a breakdown of the equipment that provides a unique assessment opportunity for each component or part ensures greater insight in determining which assemblies, subassemblies, or components significantly affect reliability and availability performance. For example, “Drivetrain-Gearbox-Bearings-Planetary Bearing” provides much more information than just “Gearbox”. Fortunately, with metadata software, this needs to be done only once for each equipment type. The elemental framework can then be copied for each turbine where only the turbine name or number is globally replaced as it is copied/duplicated.

## **1.2. Attend to the Details**

With any data collection system, one of the biggest challenges is ensuring that data is entered for every applicable data field. In addition to entering all of the relevant information, ensuring that standard and correct information is entered is also essential. There is often a trade-off to be made when weighing the value of data collected against the cost of collecting it. With the right hardware and smart software, it is possible for technicians to record data quickly and accurately without adding an unnecessary burden. A well-designed system can greatly reduce the amount of follow-up data entry and provide the quality assurance required.

## **1.3. Ease of Use**

To have an accurate, consistent, and usable system, it is important to limit the amount of time spent entering and updating records. This can be achieved by incorporating automated data collection and validation into maintenance processes. In addition to automated validation, use of handheld devices can decrease entry error and allow for automated capture of many data elements, e.g. date, time, asset, technician, etc. Typically, the manner that maintenance data will be used is not known at the time the system is implemented. Modern software systems can provide an interface that makes data entry easy and accurate and can also store information in a way that facilitates later use by the various groups who need to access the data. These software systems must allow for modification of the metadata frames that then port to all similar assets so the system can be realistic, dynamic and “future-proof”.

## **1.4. Root Cause Analysis: Down to the Bolt**

To truly understand the impact each part has on overall reliability and availability, it is important to distinguish between parts that caused a failure (primary failures), parts that failed as a result of the primary failure (secondary failures), and other parts that need to be repaired/replaced in the process of performing maintenance on parts with primary and secondary failures (ancillary failures). For example, if a power spike from a power supply causes the power supply to fail and also shorts out a circuit board under a console panel, then the power supply is a primary failure, the circuit board is a secondary failure, and the console panel is an ancillary failure. If multiple parts are worked on for the same maintenance action, a “Failed Part” field could be used to identify parts with primary failures and distinguish them from those with secondary or ancillary failures. Additionally, parts are sometimes opportunistically replaced when other maintenance events are underway, thus significantly reducing their replacement time and/or cost compared to their usual replacement time and/or cost.

#### **1.4. Root Cause Analysis: Down to the Bolt** (continued)

These opportunistic replacement activities should also be captured. In some cases, the part with a primary failure may not be obvious at the time of the maintenance event. In these cases, returning to the maintenance record after the root cause is discovered will be important to create an accurate and complete assessment of the maintenance event.

#### **1.5. Cost Tracking: Labor and Replacement Parts**

While availability and reliability are key metrics in assessing equipment performance, understanding what is driving maintenance costs can be just as valuable. Typically, the parts and personnel costs are stored outside the maintenance system and the relevant information from the maintenance system, including parts replaced and man-hours, is used to calculate the total cost for each maintenance event. This information needs to be holistically integrated with the real-time data via connectors to the enterprise resource planning (ERP) system(s). Only then can visibility from normal operation, including fault and repair, and all of the associated costs be coordinated. When this information is integrated and available to all personnel responsible for asset operation, then root cause analysis and a complete understanding of asset operation costs can be obtained and kept for future use.

#### **1.6. Part Source Identification**

For relevant event types, the source of parts should be clearly captured. This includes parts cannibalized from other equipment, purchased outside the main supply system, and acquired by other means, including parts machined on-site. Identifying the source of parts, including those exchanged between equipment, will allow for accurate cost calculations and set the stage for advanced CMMS uses, such as parts and inventory tracking. This can be accomplished through a “Parts Source” field. Integrating this information with real-time data is critical to a complete asset management system for the wind plant.

### **2. Data Collection System**

The first step to information rich decision making is accessible storage of the vast amounts of data generated by each turbine, sub-station, and ancillary equipment (Balance of Plant, or BOP).

---

## **2. Data Collection System**

(continued)

As with any storage solution, the initial step in the design process is the most important: deciding the type and use of information that needs to be extracted, analyzed, and visualized. Designing a data collection infrastructure with the end use in mind allows for the eventual retrieval, analysis, reporting, and displaying of information in a far more efficient and effective manner.

### **2.1. Design, Install, Scale, and Keep Evergreen Data Collection Systems**

Wind plant operators want tools to drive their decisions with solid data. After determining what systems are most appropriate and the general type of data that should be captured, implementation is the next step. Yet implementation is not a trivial task and requires committed and knowledgeable staff and executives both on the owner and vendor sides. Incomplete and ineffectual implementations result in high-cost systems with few benefits, often requiring replacement or upgrade before any return on investment can be achieved. Additionally, one of the most important aspects of data architecture is that it needs to evolve. Many good systems are eventually tossed aside or misused because they do not evolve and grow as the business changes. For companies that own or operate multiple turbine technologies, the initial design stage also needs to include an assessment of the data available from each technology and a way to map these data points so that equal comparisons can be made.

The design of the data collection and storage systems can be approached from two directions: internal resources and external vendor and system integrator services. Often the use of internal skills and systems facilitates a custom approach that best suits the company and its existing infrastructure. However, realistic assessment of the skills available versus the skills required is an important first step. A hybrid approach of using external contractors to fill any gaps in in-house knowledge is also an effective solution. Once it is established that the skills needed are available, a detailed cost assessment can then be performed.

Collecting data and collecting useful data are not the same, and this distinction is often the defining characteristic of a successful implementation versus an unsuccessful one. Paradoxically, the vast amount of data available can make collecting useful data more challenging. Wind plants produce staggering amounts of data. Estimated annual storage of essential supervisory control and data acquisition (SCADA) data from a plant with 100 modern wind turbines can exceed 150 gigabytes of data annually and this figure increases dramatically if every SCADA tag is stored at the highest possible frequency. Collecting too much or too little data can result in inefficient systems that do not produce the analysis results that are expected.

## **2.1. Design, Install, Scale, and Keep Evergreen Data Collection Systems**

(continued)

Decisions need to be made in advance of data collection to establish the types of analysis needed, thereby ensuring the collection of the data needed to complete the analysis. Architecting the hardware properly at the sites to minimize data issues due to undersized hardware forcing future site upgrades (hardware and software) is key also. Management of the data system is also required during implementation to tune the data streams so that unnecessary data (instrument noise) is not collected. Proper tuning can dramatically reduce disk consumption and future storage requirements. Well-managed data streams also allow rapid retrieval via the real-time data search systems.

While data collection for wind turbines is important to fully understand plant reliability, data must be collected for turbines and other equipment in the Balance of Plant (BOP). Data from meteorological towers, the substation, and the electrical collection system are absolutely necessary to understand the reliability of the turbines and the whole plant. For example, turbine availability can be 100%, but if the substation is down the plant is not producing. Failing to capture such a situation will lead to large blind spots in any reliability analysis.

## **2.2. Data Collection and Storage**

In any data storage scheme, the structure of the whole is as important as the structure of the individual parts. There are two common approaches to wind plant data storage; relational databases and data historians. Relational database (RDB) products offer storage of large data sets useful for non-real-time or instrument/equipment data. RDBs are very effective for storing asset information and other key textual data associated with the wind plant. This type of database is widely used in many implementations throughout many industries. Real-time data historian products are information technology systems that store time series data, allowing the storage of large data streams at high speed while using compression to manage the hard drive storage space needed for these millions of pieces of information. Data historians are commonly found in manufacturing, pharmaceutical, and utility industries including wind.

---

## **2.3. Supervisory Control and Data Acquisition (SCADA): Time Series Data**

One of two main types of information captured by SCADA is time series data on the turbine, BOP, and environmental conditions. For turbines, this time series data creates the “heartbeat” of the machine. It is collected almost continuously, typically once per second or more often, and is stored in regular intervals at the limit of the instrumentation and the data collection architecture. The various data streams that are captured are sometimes referred to as tags. For those familiar with traditional databases, a tag is like a database field. These data points record the operating conditions of the turbine and its parts, as well as the environmental conditions in which the turbine is operating. Many plants choose to archive their SCADA data in a real-time historian.

A multitude of time series data is available from a wind turbine SCADA system, enabling a great variety of analysis. As an example, the set of tags necessary for basic reliability analysis for a turbine is:

### **2.3.1. Turbine Status or Operating State**

Terminology can vary widely among owner/operators or original equipment manufacturers (OEMs), but some basic examples include: up and running, available but idle, down for repair, curtailed, and manually stopped at the turbine.

This value can be stored as a text field, usually with abbreviated versions of the state descriptions, or as an integer with a given number mapping to a specific description. If this value is stored as an integer, care should be taken so that it is not translated to a real number in an historian or other database. If 1 means “up and generating” and 2 means “down for maintenance”, a value of 1.62 is not very useful.

For turbine status, high-resolution data, data captured very frequently versus less often, is necessary to determine turbine status over the full course of a day, week, month, or year. When turbines are coming online and offline frequently, data that does not show the state changes do not provide enough visibility into the turbine’s true condition. This limitation can be overcome either by collecting this data at a higher frequency or by only capturing this data upon state change. Collecting upon change will yield the best data storage performance with respect to hard drive space consumption and retrieval speed.

### **2.3.2. Power Generated**

Typically stored in MW, the power generated by the turbine is very useful for reporting on turbine production. It can also be a valuable “sanity check” when various data sources are in conflict regarding the turbine’s actual status.

Most SCADA systems offer more than one power metric: turbine, string, or park. It is important to be clear which is being reported. This is managed by careful metadata and taxonomy design.

### **2.3.3. Wind Speed**

Typically, there will be at least two sources of wind speed data: the turbine’s anemometry and the meteorological tower. Both sources can be useful to understand what is really happening at a turbine.

All analyses of a wind plant’s operations need to consider wind speed. Ideally, the actual wind speed, usually measured in meters per second, should be captured. When this data is plotted in a turbine power curve, a highly effective tool for combining wind speed and power output and determining, from the shape of the curve, how a turbine is performing is available to owner/operator and OEM staff.

Beyond those listed above, many of the other turbine, BOP, and environmental tags in the SCADA time series data will be useful at some point for root cause reliability analysis. In particular, tags that are generally useful for root cause analysis include measures of temperature, including ambient air temperature and the temperature of components, measures of other air conditions, including wind speed and direction, air pressure or density, and turbulence, and vibration modes from multiple contact points and rotational speeds of the components in the turbine. Transformer gas and condition monitors are also highly useful for preventing transformer failure as replacement lead times for these components are typically several months to a year.

## 2.4. SCADA: Alarms

The second type of SCADA data relates to alarms at the turbine. Events, alarms, and faults are collected when they occur, not continuously as with the time series data, but are typically stored as a time series to correlate with the instrument and asset data streams. With this kind of data, information is only stored when something interesting happens, namely events are recorded when the operating or environmental conditions of the turbine and its parts fall outside of specific boundaries. Combined with work orders, alarm information can help provide a complete set of downtime events for each turbine, BOP equipment, and the plant. Ideally, any alarm that requires human intervention will also have a work order associated with it. As an example, turbine alarms should contain the information presented in Table A at a minimum.

*Table A*

Turbine Identifier	Recording the turbine ID links each alarm with a specific turbine.
Event Identifier	Most SCADA systems have a list of a few hundred alarm types. Capturing an identifier for each alarm, then cross-referencing the meaning from a complete list of alarms and their attributes, provides much information about what was going wrong. Attributes can include useful information, such as whether an alarm can be automatically or remotely reset and whether the alarm was triggered automatically or by human intervention. Cross-referencing can be done automatically through the construction of look-up tables and maintaining continuity between alarms, alarm codes, metadata, real-time data, and the overall plant and fleet taxonomy.
Alarm Start Date and Time	Date and time when the alarm begins.
Alarm End Date and Time	Date and time when the alarm ends.



## **2.5. Computerized Maintenance Management Systems: Work Orders**

Beyond SCADA storage, many owner/operators have also implemented computerized maintenance management systems (CMMS) for their work orders. A CMMS enables access to work order data for trend analysis, detailed parts tracking, and root cause analysis. A CMMS is a crucial, but frequently overlooked, aspect of the data collection architecture. Paper work orders and technician tribal knowledge are ineffective sources of information about turbine, BOP, and plant performance, especially over the life of the equipment and wind plant. One of the largest analysis challenges facing the wind industry is the current dependence on manual maintenance and repair documentation processes. These are not scalable and deprive owner/operators of the crucial corrective action information that is necessary for root cause analysis. Well-written work orders can provide a goldmine of information for a company while poorly-written work orders can be a waste of valuable technician time.

One of the cardinal rules of a CMMS, or any other data entry system requiring human input, is that it needs to be as painless as possible to do data entry. Automating the data collection with handheld devices, bar coding, and passive identification systems, such as radio frequency identification (RFID) can mean the difference between capturing data or missing critical pieces of the operations and maintenance (O&M) puzzle. The people involved in work order data entry can vary widely, but often include technicians, administrative staff at the plant, and employees in an operations command center (OCC). Other important aspects to keep in mind in designing data entry systems are that optional fields tend to remain blank and “miscellaneous” is a popular choice. Avoiding inaccurate and incomplete tracking and recording can mean the difference between understanding turbine, plant, and fleet performance and multiple root cause unknowns. At a minimum, high-quality work orders for a turbine should contain:

### **2.5.1. Turbine Identifier**

Recording a turbine ID links each maintenance event with a specific turbine. Events that do not tie to a specific turbine can still be captured, but this should be clearly specified. Ideally, there will be options to choose specific BOP equipment in addition to specific turbines.

### **2.5.2. Event Type**

Event type captures, at a high level, what kind of work is being performed, e.g. component failure, preventative maintenance, inspection, etc.

All downtime and maintenance events should be recorded, including inspections and other scheduled maintenance events. Even inspections and scheduled maintenance that is relatively short in duration, relatively infrequent, and/or can occur while the system is running are crucial to understanding the availability, reliability, and financial performance of a system.

### **2.5.3. Affected Component**

Ideally, the affected component would be chosen from a standard breakdown of the turbine, e.g. taxonomy, metadata framework, or equipment breakdown workflow. This value may not be initially known with certainty, so a good CMMS needs to allow for updates, editing, and refinement as more knowledge is gained.

In order to conduct real root-cause analysis, it is also useful to capture a brief description of the failure mechanism and/or the external event that caused the downtime or maintenance, e.g. curtailment, chipped gear tooth, dirty oil, etc.

For relevant event types, the source of parts is also a useful piece of information. This includes parts acquired through non-standard methods, e.g. swapped from another turbine, purchased outside the supply system, machined on site, etc. Identifying the source of parts allows for more accurate cost calculations and will allow more advanced CMMS towards parts and inventory tracking.

### **2.5.4. Equipment Status**

Not all maintenance events will stop a turbine from generating. For example, some inspections are allowed when the turbine is running.

Suggested choices for equipment status include online, offline/fault, planned maintenance, unplanned maintenance, degraded, etc. It is very important to establish categories for up and down time and for operations management to ensure accuracy and consistency amongst the engineering and technician teams.

### **2.5.5. Event Start Date and Time**

Date and time when the status of the turbine changes. Or, if the turbine status does not change due to the start of the event, the date and time the maintenance event begins.

### **2.5.6. Event End Date and Time**

Date and time when the status of the turbine changes. Or, if the turbine status does not change due to the end of the event, the date and time the maintenance event ends.

### **2.5.7. Downtime**

There are many ways to measure downtime. From the event start and end times, the total duration of the downtime or event can be captured. Other useful measures are found in Table B.

*Table B*

Active Maintenance Time	The total amount of time that maintenance was actively performed on the turbine.
Person-Hours	The total number of person-hours required to complete the maintenance action. Note that this may be very different (greater than or less than) the total downtime, and may be greater than the active maintenance time if more than one technician was needed.
Waiting Time	Ideally, this can be broken into time spent waiting for a technician to become available, waiting for a part from supply, waiting for a piece of support equipment to become available, or waiting on other administrative or supply delays.

---

### 2.5.8. Description/Comments

Though free-text comments can be difficult to use in an automated way, allowing technicians to capture anything unexpected or unusual about a maintenance event can be quite useful when delving deeply into specific events or types of events. In addition, this field can be helpful to support the collection of additional data while the CMMS is being upgraded to capture it in a more appropriate field.

## 2.6. Other Systems

In addition to the data that is captured from SCADA and work orders, supplemental turbine and plant information is also needed. ERP systems containing cost and other financial and business information provide the supplemental information needed to support data-based decision making.

Ideally, cost information would include component-level repair costs, component-level replacement costs, consumables costs, e.g. the price of a liter of gearbox oil, costs associated with technician time, and costs associated with overhead, e.g. administrative time if such overhead is linked to maintenance or downtime. Additionally, some plants also look at lost revenue from generation or penalties assessed for not generating.

Information on turbine and BOP configuration is another essential aspect of cross-fleet analysis when performing analysis at system, sub-system, component-group, and component levels, especially across multiple plants or turbine technologies. A hierarchical equipment breakdown flow or structure and the site taxonomy divide the turbines and BOP into their generalized parts in a parent-child relationship that allows sub-parts to be rolled up into sub-assemblies, sub-systems, and systems.

Once a general taxonomy is developed, then each of the turbine technologies or plants can be mapped to it, creating a standard that allows comparison. Also, a detailed description of the equipment (make, model, manufacturer of major components, presence/absence of optional systems such as de-icing equipment or condition monitoring, etc.) is important for comparisons. Lastly, documented system knowledge, such as turbine specifications or sub-station fault trees, can provide the basis for more advanced reliability analysis.

## **2.7. Data Processes**

In the wind industry, many multi-site and OEM companies have implemented operations command centers (OCCs) with real-time operating data flowing from plants to a centralized monitoring and control center. The real-time data is then stored in a single large enterprise-level database covering multiple plants. This approach requires a robust and reliable connection from each turbine and plant to the OCC or reliable data storage at each site so that when connection to the OCC is restored the buffered data is passed to the OCC. An alternative, seen at smaller operators and plants, is where a storage system is implemented on-site and stores a finite time period of SCADA data. Some of these implementations allow for subsets of the data to be sent to a central office for storage and analysis periodically or after events.

Those companies that do store their SCADA data consider it highly proprietary and treat it as intellectual property. This adds a requirement for encryption and security during the transfer of data from the plant and access levels and controls that restrict who can view the stored data. If there is transfer of the data from a plant to an OCC, a hardened high-bandwidth connection is most desirable. This creates a dedicated connection between the wind plant and the OCC, making it a good choice for carrying large amounts of data, as it is both reliable and secure. Once data is stored at the OCC, the use of integrated security protocols can fulfill the needs for controlling access to the data.

After data is stored and accessible to those with the rights to see and use it, data protection becomes a primary task of the data administration staff. Design, implementation, and maintenance of a backup and recovery plan are essential to preventing the loss of data through accident, data corruption, hardware failure, or natural disaster. The plan should include levels of criticality for the data, projected recovery timeframes, scheduling and monitoring of backups, on-going validation testing of the backups, and the media choices on which the backups will be stored.

In addition to backups with the amount of data being stored for each turbine and plant, an archiving strategy is necessary to manage the size of the database and maintain a high-functioning retrieval system. One approach is to archive the raw data but to retain calculated values.

## 2.7. Data Processes

(continued)

Another approach is to store the data using special compression techniques that reduce the amount of data stored without losing the meaning of the data. Data historians are specially designed with this type of compression in mind. With the cost of storage coming down over the prior decades, archiving strategies should be periodically re-evaluated to ensure that the correct levels of data are available for analysis.

When setting up transfer and storage protocols, the data to be stored must be determined. This concern is especially relevant when looking at the need to summarize the voluminous SCADA time series data. For monthly or yearly performance metrics balanced against the detail needed for root cause analysis, real-time data compression algorithms become critical to balance data storage needs against data completeness. Compression means reducing the number of electronic bits that represent a piece of data, thus reducing the number of bits that need to be transferred from the plant or stored. For example, only storing values when they change can save a great deal of space if that data does not change often.

One of the other aspects of data integrity is addressing missing or illogical data, with data validation serving as an essential aspect of any data collection system. When there is only a single piece missing, it will likely have little to no impact on analysis, but when larger amounts of data are missing, perhaps covering hours, the loss may be important. The practice of data editing or filling in the data with realistic values can assist in creating a complete data set. For illogical data, values for a piece of data can be compared to previous values or sets/ranges of acceptable values, allowing an unrealistic value to be identified. Care must be taken with data editing, as it can reduce confidence in the data as a whole. Among other challenges, important signals can be missed if unexpected, but accurate, values are overwritten. Also, filling in unknown values can mask a data communications problem.

## **2.7. Data Processes**

(continued)

For all of these data integrity concerns, their impact can be reduced by implementing good business processes and procedures, where all employees follow the same process when dealing with the data. Whether the employee is at the plant, in the corporate IT department, or in the engineering/analysis group, business processes allow for the same methodology to be implemented and for necessary improvements to be implemented systematically. These remedies should include a standard approach to the use and interpretation of data. This creates an environment where comparisons between turbines and plants can easily be made because the analysis is based on the same assumptions about the data.

## **2.8. Integrate Data**

One of the greatest challenges in using CMMS and SCADA data to perform reliability analysis is in matching work orders, SCADA time series, and SCADA alarm data. This linking of symptom, e.g. high SCADA temperature recordings followed by a gearbox over-temperature alarm, to corrective action, e.g. a work order to replace a lubrication oil pump, allows for the beginning stages of root cause analysis, parts tracking, and trending. An automated method for performing this linking will greatly improve the detail and accuracy of reliability analysis, but it is not an easy process. Challenges in linking data can include conflicts between CMMS and SCADA regarding turbine status, incomplete work orders, and missing SCADA data. Additionally, real situations that are difficult to interpret will appear, such as curtailment, overlapping work orders, and back-to-back alarms. While no systems emerge as the complete solution after a plant or fleet reaches commercial operation date (COD), continuous investment and improvements to the operational and maintenance systems are crucial for assuring long asset life and the highest levels of production output.

## **3. Analysis**

The culmination of the above two sections is analysis. Analysis incorporates the data generated, collected, and made available to improve the understanding of the current reliability and performance of the wind plant and its turbines. A staged approach is required to impact an O&M strategy through improved availability, increased reliability, and reduced O&M costs. First establish baseline performance to understand what the current situation is, then identify performance drivers and determine their root causes, and finally create action plans for addressing those drivers with higher impact.

### **3.1. Understand Current Performance**

Successfully answering questions, such as “What is the current performance?” and “How good is it?”, is the first step to making improvements. This will point toward problematic areas on which to focus. Examples of questions and analysis related to baseline performance include:

#### **3.1.1. What is the Baseline Performance?**

- Calculate basic operations and reliability metrics, such as availability, MTBE (mean time between events), mean down-time, and capacity factor for the plant and then each turbine.

#### **3.1.2. How Does the Plant Performance Compare to the OEM or Financial Expectations?**

- Identify and graph how a typical turbine spends its time (what percent of the time is it running, idle and available, down for scheduled maintenance, curtailed, etc.).
- Be sure to identify when the turbine state cannot be determined, such as when SCADA communication is lost or the historian briefly stops recording.

#### **3.1.3. Are the Data Aspects of the Operations and Maintenance Processes Well Understood?**

- Make and document assumptions about the data being gathered and how it is gathered, stored, and used for analysis. Institutionalize these assumptions so that all departments have the same meaning for particular pieces of data.

### **3.2. Identify Performance Drivers**

Once baseline performance is understood, then performance drivers can be found. Methods for identifying these drivers can include exploring trends, outliers, good performance, and surprising results. Examples of questions that can be answered at this point and their related analyses include:



### **3.2.1. What is Driving Poor Performance?**

- Identify key contributors to low generation, unavailability (downtime), etc. Exploring top contributors is a simple, but very powerful method for identifying areas for improvement.
- Compare multiple metrics, such as event frequency versus event duration or generation versus turbine wind speed. The outliers are especially interesting in these types of graphs.

### **3.2.2. Where is Performance Roughly the Same? Where is There Great Variability?**

- Explore turbine-to-turbine performance and variability in all the basic aspects including availability, MTBE, mean downtime, and capacity factor.
- Explore daily, weekly, monthly, and seasonal trends. Plot graphs of the metrics over time. Look at the whole plant and also look at individual turbines or individual event types, especially those with very high or very low performance.

### **3.2.3. Where Are the Business Data Processes Different?**

- Address any inconsistencies in data processes and assumptions, including determining if there is a valid reason for doing things differently.
- Understand limitations in the data systems, analysis/modeling, and reporting.

## **3.3. Determine Root Cause**

After understanding baseline performance and identifying some of the key performance drivers, then root causes can be identified to solve problems. Examples of questions and analysis that can be addressed at this point include:

**3.3.1. Why are Certain Aspects of Operations, e.g. Turbines or Groups of Turbines, Months or Days of the Week, Types of Scheduled Maintenance, Having Such a Negative Impact?**

- Investigate de-rates and periods of unexplained performance.
- Interpret unexpected patterns.

**3.3.2. What are the Root Causes of the Top Problems?**

- After identifying the operational aspects that have the most impact, conduct root cause investigations. Follow through on this activity. Simply identifying potential root causes is not enough. Real fixes should be developed, tested, implemented, and assessed.

**References**

- [1] V. A. Peters, P. S. Veers, and A. Ogilvie, "Wind Energy Computerized Maintenance Management System (CMMS): Data Collection Recommendations for Reliability Analysis," Sandia, Albuquerque, NM, USA, SAND2009-4184, 2009.
- [2] B. L. McKenney, A. B. Ogilvie, and V. A. Peters, "Using Wind Plant Data to Increase Reliability", Sandia, Albuquerque, NM, USA, SAND2010-8800, 2011.